7В

УДК 004.89:004.4

А.С. Вороной
Государственный университет информатики и искусственного интеллекта,
г. Донецк, Украина
postmaster@iai.donetsk.ua

Виртуальная организация агентов для разработки онтологической базы знаний системы компьютерного обучения

В работе рассмотрена организация мультиагентной системы построения онтологий предметных областей для целей компьютерного обучения. Предложен состав виртуальной организации агентов для автоматизации поиска информационных ресурсов, анализа текстов с целью выявления концептов и отношений между ними, разработки и пополнения онтологий предметных областей, включения созданных онтологий в репозиторий знаний.

Введение

Современные модели представления и интеграции информационных ресурсов активно развиваются и внедряются в практику электронного обучения. Для работы с такими ресурсами, их эффективного поиска в соответствии с образовательными программами необходима возможность описания характеристик и семантики ресурсов и предметной области, а также возможность обработки данной информации программными агентами. Важнейшим элементом современных информационных технологий являются онтологии, которые позволяют производить автоматизированную обработку семантики информации. В основе онтологий лежат свойства, классы, объекты и ограничения, реализующие представление об объектах как о множестве сущностей, характеризуемых некоторым набором свойств. Эти сущности состоят между собой в определенных отношениях и объединяются по определенным признакам в классы. В результате полного описания объектов и их свойств предметная область представляется как сложная иерархическая база знаний, над которой можно осуществлять «интеллектуальные» операции для создания персонального контекста обучения и генерации контрольных заданий для тестирования.

Для решения задачи создания онтологической базы знаний для целей компьютерного обучения представляется целесообразным использование многоагентных систем (МАС) и технологий Semantic Web. В области исследований и разработок МАС в последние годы активно проектируются виртуальные организации агентов, которые характеризуются решаемыми задачами, определяются составом и функциями агентов, распределением их ролей и полномочий, способами коммуникации.

В **данной работе рассматривается** решение задачи разработки состава и определения функций агентов виртуальной организации, автоматизирующей построение онтологической базы знаний для системы компьютерного обучения. Работа виртуальной организации основана на кооперации рабочих групп агентов, решающих задачи этапов: создания, пополнения, интеграции онтологий предметных областей для онтологической базы знаний системы компьютерного обучения.

Состав и функции агентов системы автоматизации разработки онтологий

Разработка онтологий предметных областей (ПрО) знаний, используемых для целей обучения, является достаточно трудоемким и длительным процессом, требует участия коллективов экспертов. Поэтому актуальной является задача автоматизации процессов построения онтологий. Все этапы разработки онтологий должны поддерживаться всевозможными средствами автоматизации процессов: поиска текстовых документов с информацией по предметной области, преобразования текстов на основе формализмов представления смысла, анализа смысла текстов с целью выявления концептов и отношений между ними, построения онтологии и представления ее в виде, удобном для хранения.

Для сокращения затрат времени на создание онтологий ПрО в системе автоматизации создания онтологий будем использовать следующие подходы:

– использование базовых тезаурусов или базовых онтологий для предметной области;

– использование шаблонов онтологий для групп предметных областей;

– использование существующих частичных онтологий для разделов ПрО.

Поскольку работа по созданию онтологии должна быть коллективной, систему автоматизации процесса разработки онтологий целесообразно строить на основе многоагентного подхода. В состав виртуальной организации агентов системы создания онтологий должны входить следующие рабочие группы (РГ) агентов:

– поиска документов по ПрО в неструктурированных источниках и семантическом Web(АПД);

– извлечения знаний из документов и представления их в формализованном виде, удобном для включения в онтологию (АИЗ);

– интеграции онтологий и объединения фрагментов (АОФ);

– обеспечения хранения и доступа к онтологии и ее фрагментам (АОД).

Исходными данными для разработки онтологии ПрО являются словарь или базовый тезаурус предметной области и онтологии верхнего уровня (описывают наиболее общие концепты: пространство, время, материя, объект, событие, действие и т.д.)

Рассмотрим состав и роли агентов рабочих групп.

В состав рабочей группы АПД входят агенты поиска, агенты проверки существования документов, агенты ранжирования документов, агенты получения документов. Координацию работы агентов выполняет агент-менеджер рабочей группы АПД. Поисковые агенты используют поисковые системы тематических коллекций по ПрО, многоцелевые поисковые системы, базы данных «скрытого» Web, системы поиска семантического Web. Агенты проверки существования документов устанавливают актуальность ссылок на ресурсы, выданных агентами поиска. Агенты ранжирования документов исключают повторяющиеся ссылки и упорядочивают по степени релевантности ссылки на все документы, найденные поисковыми системами. Степень релевантности вычисляется на основе аннотаций, выдаваемых поисковыми системами. Менеджер РГ АПД на основе ранжированного списка ссылок формирует задания агентам получения документов.

Наиболее интеллектуальными в виртуальной организации являются агенты извлечения знаний. Документы, полученные РГ АПД, могут быть двух видов: тексты на естественном языке (ЕЯ-документы), документы семантического Web с формализо-

ванным представлением семантики (ФС-документы). Для обработки документов двух классов в РГ АИЗ предусматриваются, соответственно, агенты формализации смысла документа (АФС) и извлечения знаний из ЕЯ-документов для включения в онтологию (АИЗ ЕЯД), а также агенты извлечения знаний из ФС-документов для пополнения онтологии ПрО (АИЗ ФСД). В качестве формализма для представления смысла текста в ЕЯ-документах предполагается использовать семантические Q-сети [1] на основе аппарата пирамидальных сетей (ПС) В.П. Гладуна [2] и семантические представления И.П. Кузнецова [3]. Элементы ПрО описываются в ЕЯ-текстах элементарными и составными словосочетаниями. Первые часто состоят из двух слов и являются реализациями элементарных отношений (свойство – объект). Вторые можно представить в виде комбинации элементарных словосочетаний. Они являются реализациями составных отношений. Понятия ПрО также выражаются словосочетаниями (одним словом – наименованием понятия, элементарным словосочетанием или комбинацией элементарных словосочетаний). На основе анализа семантического представления текстов в виде Q-сети и использования базового тезауруса ПрО определяются отношения для включения в онтологию. Использование словаря ПрО позволяет отфильтровать из общего набора полученных семантических конструкций лишь те, которые играют существенную роль в онтологической картине предметной области.

Работа АИЗ ФСД определяется языковыми средствами представления семантики документа (RDF, OWL) и способом представления фрагмента онтологии.

Группа АОФ состоит из агентов включения фрагментов (АВФ) и агентов интеграции (АИ) вариантов онтологии для одной ПрО. АВФ добавляют фрагменты, полученные агентами извлечения знаний, в создаваемую онтологию. При этом устанавливается уровень концепта в иерархии «класс – подкласс» и создаются требуемые связи в графе. АИ решают задачи сравнения онтологий, выравнивания порядков онтологий, объединения классов и их свойств и отношений, разрешения конфликта имен.

Группа АОД содержит агентов записи (АЗ) созданной версии онтологии в базу знаний и агентов выдачи (АВ) фрагмента онтологии по запросам интерфейсных агентов пользователей-экспертов (АПЭ).

Рассмотрим коммуникации рабочих групп агентов при обработке запросов от агента пользователя-эксперта для двух случаев: при создании онтологии для новой предметной области (задача 1) и при пополнении существующей онтологии новыми знаниями (задача 2).

В первом случае пользователь-эксперт в запросе к рабочей группе АПД указывает название ПрО. Менеджер РГ АПД формирует задание поисковым агентам на основе онтологии верхнего уровня и базовой онтологии ПрО, если она имеется, или на основе базового тезауруса. Релевантные результаты поиска, подтвержденные экспертом, в виде текстовых документов передаются менеджером РГ АПД менеджеру РГ АИЗ. В данной рабочей группе документы классифицируются менеджером и передаются на обработку соответствующим агентам: АИС ЕЯД и АИС ФСД. Результатом работы данной группы агентов будут фрагменты онтологии ПрО для обнаруженных концептов. Полученные фрагменты группа АОФ объединяет в онтологию ПрО и передает ее агентам АОД для помещения в базу знаний.

Во втором случае существующая онтология используется для формирования поисковых запросов по предметной области и поиска документов с новыми знаниями. При формировании запросов при этом используются также шаблоны для групп предметных областей. Найденные релевантные документы используются для поиска новых концептов.

Диаграмма коммуникаций рабочих групп агентов при решении задач 1 и 2 показана на рис. 1.

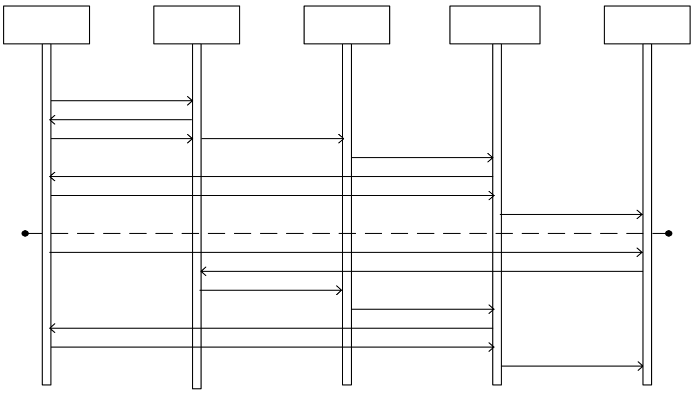


Рисунок 1 – Диаграмма коммуникации рабочих групп агентов

Выводы

АПЭ РГ АПД Р

Предложенная мультиагентная система автоматизации разработки онтологий реализуется в проекте интеллектуальной обучающей среды для цикла дисциплин направления подготовки студентов «Компьютерные науки» Государственного университета информатики и искусственного интеллекта. Применение разрабатываемой системы позволит сократить время создания онтологической базы знаний.

Литература

1. Формирование базы лексических функций и других отношений для онтологии предметной области / Н.Г. Загоруйко, А.М. Налетов, А.А. Соколова, В.А. Чурикова // Труды Международной конференции Диалог-2004. – М. : Наука, 2004. – С. 202-204.
2. Гладун В.П. Планирование решений / Гладун В.П. – Киев : Наукова думка, 1987. – С. 17-51.
3. Кузнецов И.П. Семантические представления / Кузнецов И.П. – М. : Наука, 1986.

О.С. Вороной

Віртуальна організація агентів для розробки онтологічної бази знань системи комп'ютерного навчання

У статті розглядається організація мультиагентної системи побудови онтологій предметних областей для комп'ютерного навчання. Запропоновано склад віртуальної організації агентів для автоматизації пошуку інформаційних ресурсів, аналізу текстів з метою виявлення концептів і відношень між ними, розробки і поповнення онтологій предметних областей, включення створених онтологій у репозиторій знань.

A.S. Voronoy

Virtual Organization of Agents for the Development of Ontological Knowledge Base for Computer Learning System

In this article organization of multiagent system for constructing ontologies for computer learning is proposed. The contents of virtual organization of agents for automatic search of infomational resources, text analysis with the goal of detecting concepts and their relations, development and adding ontologies, adding creating ontologies in knowledge repositories is proposed.

Статья поступила в редакцию 13.07.2010.